

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

# MPEG-4: Context and Objectives

**Rob Koenen**

KPN Research  
St. Paulusstraat, 4, 2264 XZ Leidschendam, The Netherlands  
E-mail: [r.h.koenen@research.kpn.com](mailto:r.h.koenen@research.kpn.com)

**Fernando Pereira**

Instituto Superior Técnico / Instituto de Telecomunicações  
Av. Rovisco Pais, 1096 Lisboa Codex, Portugal  
E-mail: [fernando.pereira@lx.it.pt](mailto:fernando.pereira@lx.it.pt)

**Leonardo Chiariglione**

CSELT  
Via G. Reiss Romoli, 274, 1048 Torino, Italy  
E-mail: [leonardo.chiariglione@cse.lt.stet.it](mailto:leonardo.chiariglione@cse.lt.stet.it)

## Abstract

*The MPEG-4 audiovisual representation standard currently under development addresses the needs and requirements, arising from the increasing availability of audiovisual content in digital form. It goes beyond digitizing linear audio and video, specifying a description of digital material in the form of 'objects', that can be flexibly and interactively used and re-used. This paper describes the developments that make MPEG-4 necessary ('Why'), what MPEG-4 does to address the new needs ('What'), and gives an overview of how the standard is developed ('How').*

## 1. Introduction

After setting the MPEG-1 and MPEG-2 standards, MPEG (Moving Pictures Experts Group, ISO/IEC Joint Technical Committee 1, Sub Committee 29, Work Group 11) is now working on a new audiovisual standard, called MPEG-4. Many people who have heard of MPEG-4 still seem to think that it is about 'very low bitrates'. Although, at the beginning of the work on MPEG-4, this was indeed the objective of the new standard, MPEG adapted the work plan to the changes in the audiovisual environment and modified its targets considerably. The new MPEG standard under development now has a much wider purpose: it addresses the new demands that arise in a world in which more and more audiovisual material is exchanged in digital form. These needs go much further than achieving even more compression and even lower bitrates.

The first two sets of MPEG standards (MPEG-1 and MPEG-2) are well-known to people involved in digital communication. They are widely adopted in commercial products, such as CD-interactive, digital audio broadcasting, digital television and many video-on-demand trials. MPEG-1 and -2 deal with 'frame-based video' and audio: although these standards provide a large improvement, in randomly accessing content, over standards that existed before, the granularity of the interaction is limited to the video frame, with its associated audio. In this sense, the functionality could be compared with that of audio and video cassette players, albeit with non-linear controls. Their most important goal is to make storage and transmission more efficient, by compressing the material. The new MPEG-4 standard does not only aim to achieve efficient storage and transmission, but also to satisfy other needs of (future) image communication users. To reach this goal, MPEG-4 will be fundamentally different in nature from its predecessors, as it makes the move towards representing the scene as a composition of (potentially meaningful) objects, rather than 'just' the pixels.

This paper will explain what MPEG-4 is trying to achieve, as well as why and when. The 'how' of MPEG-4 is only briefly touched upon; the other papers in this special issue of Image Communication go into that question with much more detail.

## 1. Why: Relevant Developments, But Not Yet Supported

Digital audiovisual systems for consumers are being deployed rapidly. Until very recently, they all had a number of characteristics in common: they were predominantly hardware-based - meaning that the decoding of the audiovisual material is carried out by dedicated chips - the coding was ignorant of the coded material's meaning, and the user had only a very limited amount of control over the presentation and the configuration of the system.

The new MPEG-4 standard is necessary because the ways in which audiovisual material is produced, delivered and consumed are still evolving. Furthermore, hardware and software keep getting more powerful. In the following paragraphs, these changes are elaborated.

The **production** of audiovisual material is changing: the only way to produce content used to be recording it, with a camera and microphone. In recent years, however, an increasing part of audiovisual content is computer-generated. Artificially generated content used to be limited to movies and science, but it is now quite common to see it used in other business applications -for example architecture-, and consumer devices, such as computer-games. This computer-generated type of content is often 3-dimensional, and also naturally captured scenes in which depth plays a role are starting to appear. Next, it is very important to observe that natural content is more often produced and stored as 'pre-segmented' material, usually using blue-screen techniques. Lastly, we note that during the production and subsequent organization and storage of audiovisual content, information of various natures -notably textual information- can be associated with the content for later use.

The **delivery** of multimedia information is changing: new networks carry video and other data, next to the audio information that they were originally intended to transport. This applies to the PSTN, but also to mobile networks. Users will want to have, when they are traveling, the same possibilities that they enjoy at home or in their offices - which includes access to multimedia information. Even the GSM network, with net data rates below 10 kilobits per second, is starting to be used for transmission of digitally compressed video. On the other hand, high bandwidth networks are also becoming available. Cable modem and ADSL technology both promise to deliver several Megabits per second to users, without the need to change access network cabling. Further, as a larger variety of networks carry multimedia information, the number of connections that involve more than one network type is growing.

Not only are the production and delivery of multimedia information changing, **consumption** is too. First, more and more content is audiovisual, which means that multimedia information is finding its own place next to the common textual data. Second, a growing part of the information is read, seen and heard in interactive ways. The most obvious environments in which this can be observed, are the personal computer and the Internet, or more specifically, the World Wide Web.

This development in personal computing highlights another evolution as well: multimedia communications are moving away from dedicated signal processing chips, to make room for **software implementations** running on powerful generic hardware. The recently finalized ITU recommendation H.324 ('Terminal For Low Bitrate Multimedia Communication') [2] was explicitly designed to also run in 'software-only' mode, that is, on generic processors, such as those employed in personal computers. And indeed it is expected that in 1997 millions of computers will be sold with H.324 software already installed.

Another relevant development that must be mentioned here concerns **re-using audiovisual material**. The digital form of storage and transmission offers the possibility to copy audiovisual scenes as many times as one would like, without the loss of quality inherent to analog re-use. But the concept of 're-use' goes further than just the content. It applies to the tools as well. Audiovisual tools made to run in software in one box (e.g. a PC or a Set Top Unit), are in principle also usable in a different program running on the same box. They could even be used in another box, certainly if the latter has a similar processor - but also when it does not, although that is a considerably more difficult case to solve.

The last issue to note here is the **blurring of the borders between three traditionally distinct service models**: 'communication', 'interactive' and 'broadcast'. These service models can roughly be mapped on the Telecommunication, Information Technology (IT) and Entertainment sectors, respectively. It is often argued that these sectors are merging. The authors do not believe that there is really a convergence of *businesses*, but we do, however, envision the disappearance of the clear distinction between the different service models. In fact, this is already happening. Two examples are: interactivity is added to broadcast services, and mixed communication/interactive applications are appearing on the Internet. The problem, however, is that each of these sectors, in crossing the borders, brings its own technological solutions, and as a consequence, different solutions will appear for the same types of applications, which hampers interworking and common solutions. To give an answer to this development, MPEG-4 should be a standard that supports all of these service models and, by doing so, also allows mixed models to be built.

## 1. What: Coding of Audiovisual Objects

Currently available audiovisual coding systems and standards do not address the developments mentioned in the previous section, or address them only partly. Interactivity is limited to the temporal aspect of content: playing in two directions at variable speed - and even this usually only showing video. The same applies to re-use: this is limited to taking a linear piece of rectangular video and its associated audio, and inserting that in another application. Integration of natural and synthetic content in the same scene is difficult. Accessing multimedia information on new, often error-prone networks is not well-supported, nor is access and transmission across heterogeneous networks. Lastly, existing standards do not fully take the rapid progress in hardware and software technology into account.

These observations directly lead us to the goals of the new MPEG-4 standard. It is useful to first mention the most important innovation that MPEG-4 brings: it defines an audiovisual scene as a coded representation of 'audiovisual objects' that have certain relations in space and time, rather than 'video frames with associated audio' (We will refer to the old representation as 'frame-based video'). Such an object could be a video object: a car, a dog, or the complete background. It could also be an audio object: one instrument in an orchestra, the barking of the dog, a voice. When an audio and a video object are associated, an audiovisual object results: the image of a running dog together with the sound it makes. This new approach to information representation allows for much more interactivity, for versatile re-use of data, and for intelligent schemes to manage bandwidth, processing resources (e.g. memory, computing power) and error protection. It also eases the integration

of natural and synthetic audio and video material, as well as other data types, such as text overlays and graphics.

### 1. Offering a New Kind of Interactivity

Separately coding the individual objects in a scene is a very powerful tool that allows the removal of a number of the limitations mentioned above. First, it enables interaction with meaningful objects within the scene, a kind of interactivity extending well beyond the 'video player'-type of activity that analog and digital systems offer today. It will, for instance, allow the connection of information or an action to an object in a scene. One might want to associate a Uniform Resource Locator (URL) to a person appearing in a stored sequence; the URL could point to that person's home page. Clicking on the person would activate the URL.

This would be difficult to manage with frame-based video, since unless the person has a fixed, pre-defined position, there is no simple, generic way to determine if the click occurred over a particular object. Also, imagine the possibilities of re-using data once the possibility exists to separately store and access objects, rather than frames. It will be easy to create your own content, by combining several of these stored objects.

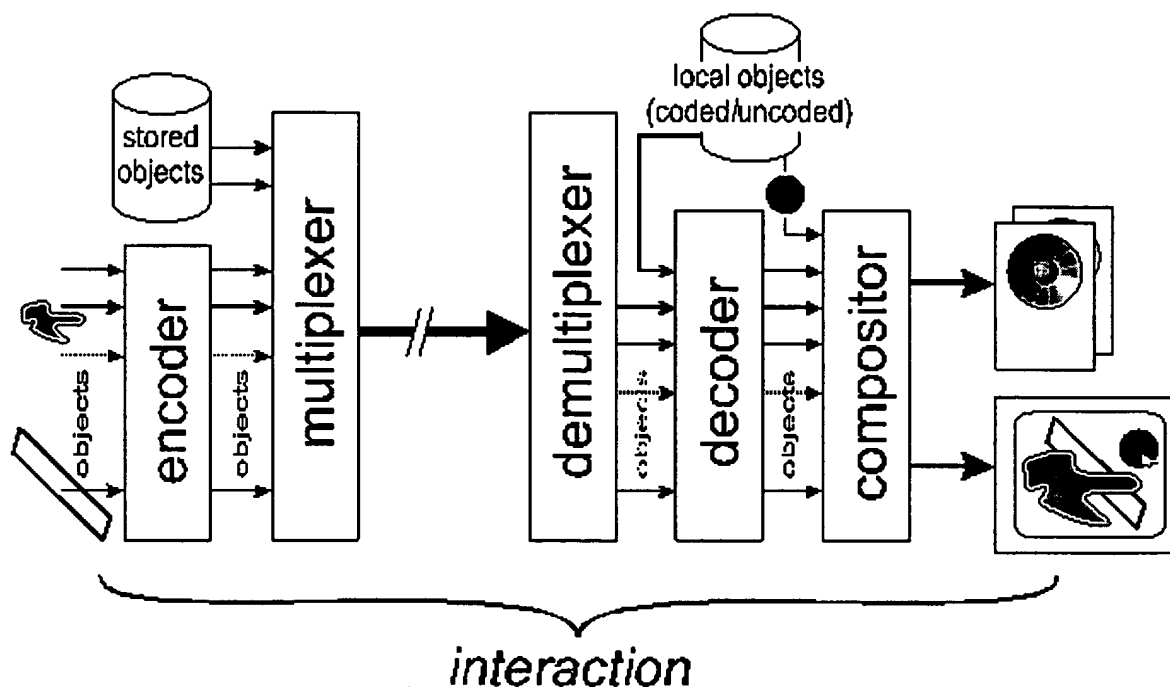


Figure 1 - Schematic overview of an MPEG-4 System

Next to the object-based representation, there is another key concept in interactivity and re-use: the introduction of a 'compositor', that takes care of the presentation of the objects. (Figure 1 gives a very general overview of an MPEG-4 system. It does not claim completeness, but shows the place and function of the compositor.) The objects that make up a scene are sent or stored together with information about their spatio-temporal relationships, or, in other words, composition information. The compositor uses this information to reconstruct the complete scene again. Composition information is thus used to synchronize different objects in time, and to give them the right position in space. Separating this function from the pure decoding of objects introduces the possibility to influence the presentation of a scene, on a screen and through loudspeakers. The easiest example: you can decide to drag an object to another place - and you can hear its associated sound move as well. You might also want to change the speed of a moving object in the same scene, or make it rotate. The possibilities of interaction, however, are not limited to the composition. Coding the individual objects separately also makes it possible to influence which objects are sent, and with which quality and error protection. In addition to that, it permits compositing a scene with objects that arrive from different locations.

It should be noted that the MPEG-4 standard will not prescribe *how* any given scene is to be organized in objects (segmented). The reason for this is that MPEG-4, just like its predecessors and other image coding standards, is a *decoding* standard, that does not specify the encoding. Actually, the segmentation is usually regarded to take place even before the encoding, as a pre-processing step, which is never a standardization issue. By not specifying pre-processing and encoding, the standard leaves room for manufacturers of systems to distinguish themselves from their competitors, by providing better quality or more options. It also allows the use of different encoding strategies for different applications. Lastly, it leaves room for technological progress in analysis and encoding strategies.

But let us, for the sake of the argument, assume for a moment that MPEG *would* specify the segmentation of a scene. Then MPEG would have to make assumptions about the use of that segmentation, and about the interpretation of what is in the

scene. This, however, will depend on the application. Rather than following this road, MPEG-4 gives the tools to represent *any* kind of segmentation, according to the principle of 'specifying the minimum' to allow maximum usability.

We find that many people ask: *"How can MPEG-4 be useful if the image analysis problem is not completely solved?"* The answer is threefold. Firstly, some problems *can* already be tackled, even in real-time (e.g. finding and tracking a face in a videophone scene). Secondly, much material, notably for interactive applications, is produced off-line, in such a way that segmentation is already available, or 'almost' available (in the case of a scene shot against a blue screen). An object-based representation does, therefore, not always require the real-time, fully automated segmentation of generic content. Thirdly, the problems that have not been solved today may well be solved tomorrow as research is ongoing; leaving analysis out of the specification allows the standard to be even more useful when this happens.

Finally, and without going into details, the authors would like to note that creating the possibility of re-using audiovisual objects over and over again, implies that copyright is an issue to take care of within the framework of the standard. Objects should have the option to carry copyright information, e.g. stating that re-use is not allowed, or giving a pointer to the owner or caretaker of the copyright.

## 1. Accessing Multimedia Information Everywhere

The standards that were set for multimedia communication so far, more or less assume an error free channel. Although H.261 and MPEG-2, for instance, have some capabilities to detect and even correct errors, they are not suitable for real time and low delay transmission of video information over networks with severe errors, such as a GSM channel. MPEG-4 provides coding tools and a systems layer that can cope with these errors. The concept of coding objects is also helpful in this case: important objects can be better protected against errors. In addition to that, the user, or the encoder itself, can decide that the bits should be spent on the important parts of a scene, which is useful as mobile networks are often of low bandwidth. This requires interaction with the encoding process.

Although this will be a difficult goal to accomplish, MPEG-4 also wants to make communication across heterogeneous networks possible. A key tool to allow this is scalable coding. A scalable object stored on a database can then be accessed at different bitrates, offering the user the best quality possible for the available network. This 'universal accessibility' principle explains why low bitrate applications, while not the prime focus, are very important to MPEG-4 - they will benefit considerably from the features that the new coding approach offers.

Accessing multimedia information 'everywhere' does not only imply supporting different and often difficult networks, it also involves allowing MPEG-4 to run on different types of machines, under different conditions. Also here scalability is important; this time it is not for accommodating different transmission speeds, but for allowing decoding on processors with varying processing power. If the scene is too complicated for the processor, this processor can decode at a lower quality, again using only part of the bitstream. An object based representation is once more an advantage, because the important objects can be given precedence in decoding. The user (or the application itself) could decide that the background of the scene is not so important, and can be decoded at a lower spatial resolution, or at a lower frame rate, or only once, or possibly not at all. The same applies to audio objects, which can be decoded at higher or lower quality, or at different bandwidths.

## 1. Integration of Objects of Different Nature

The object-based representation, that is at the heart of MPEG-4, allows a good integration of audiovisual material of different origin. The objects could be captured with a microphone and camera, ('natural'), or generated by a computer (synthetic). They could also be graphics or text overlays. Both natural and synthetic video objects can be either 'flat' (2 dimensional) or have a third spatial dimension. They can also have multiple viewpoints, which is for example used in stereo vision: these objects are generally referred to as 'multiview' objects. Different types of objects can have their own efficient representation. Each type of object is decoded by its own decoding tool, yielding an object represented in local time and space coordinates; then the compositor performs the rendering and composition of the elementary objects in the scene, using the description of the scene structure and the Transform and Properties attached to the objects. The output of the compositor is stored in an audio frame (Waveform) and a video frame (Image). A module called the 'presenter' performs the display and event handling of the scene.

A further aspect of integration is creating an audiovisual scene with objects from different sources. An example: while editing a multimedia presentation, e.g. for educational or professional purposes, the content is gathered from several databases and integrated at the user's premises. Another, more complex example: in an interactive game, local objects (representing a player, or the scenery) are combined with objects coming from remote sources (e.g. representing other players). This example nicely highlights yet another side of integration: combining real-time information (the player objects) with stored information (the scenery). Note that all the objects involved in the example - scenery and players - could be of both synthetic and natural origin!

## 1. Flexibility for a Fast-Changing Environment

As was established in Section 2, audiovisual systems that used to run exclusively on dedicated chips are moving to (also) run on generic hardware, like personal computer platforms. In principle, this allows much more flexibility than previously

possible. And indeed such flexibility would be useful in audiovisual communications to improve composition capabilities, to adapt the configuration of the decoder to the application, or to optimize it for the specific decoding platform. There are more reasons for wanting flexibility: as hardware and software technology are still progressing rapidly, it is conceivable that improvements in coding technology could relatively quickly cause parts of a standard like MPEG-4 to become outdated. This calls for the standard to be organized as a flexible toolbox. It would then be possible to customize terminals in terms of composition capabilities and decoding configuration to suit specific needs, and the toolbox can be augmented as technology progresses, without rendering the entire system obsolete.

However, when it comes to implementing a system, flexibility will not always be needed. In some cases, it will be more useful to have a highly optimized, efficient, application-specific system, rather than a program running on a general purpose chip. A clear example of such a situation is a mobile database retrieval terminal. It must be small, lightweight and highly efficient in terms of energy consumption. Resources are not to be wasted here.

The flexibility in MPEG-4 is to a large extent realized at the systems layer, also known as the MPEG-4 Systems and Description Languages (MSDL). To allow highly optimized, fixed systems on the one hand, and flexibly programmable systems on the other, different types of flexibility will have to be distinguished, taking into account the current technological limitations and the need to guarantee interworking between MPEG-4 terminals.

When MPEG-4 becomes a standard in 1998, the state of technology -read: the power of virtual machines- will allow the flexible composition of audiovisual information. It is still unclear whether real-time performance (and hence interworking) can be guaranteed in the case of downloading the configuration of tools or (more demanding) the complete tools themselves. At the same time, composition is the area where the market has shown an increasing interest and composition is therefore the current focus of the work on flexibility in MPEG-4. At the moment technology allows it, and market needs demand so, work on flexibility in MPEG-4 will be augmented, to also cover the configuration and downloading of tools 3. Although this type of flexibility is not actively pursued at the moment, in principle, the object-oriented architecture adopted in MPEG-4 already allows for it.

In terms of MPEG-4 architecture this means that two types of systems will exist, depending on receiver programmability: 'fixed', where no programmability is possible, and 'flexible' which for the moment means that composition flexibility is supported. Systems with different flexibility will be able to communicate with each other. This kind of interworking is very much an issue of 'profiles', which will be explained below. In this context, suffices it to say that a flexible system can always configure itself to act as a fixed one.

Summarizing, flexibility will start with composition and will, when compute power allows this and a market demand exists, be extended to decoder configuration and decoding tools downloading. Note that also allowing the re-use of data objects is a form of flexibility that MPEG-4 will offer, although this does not strictly need the programmable MPEG-4 system that was defined above.

## 1. How: Development and Structure of the Standard

The MPEG-4 standard is developed in a way similar to its predecessors, which means that the development process has a competitive phase followed by a collaborative phase. The competitive phase started with a widely distributed Call for Proposals 4, accompanied by a Proposal Package Description [5]. These were first issued in November 1994, at the 28th MPEG meeting in Singapore. The Call for Proposals was refined during the following meetings, asking for technology that could fit the objectives described in Section 3. The Call for Proposals and the Proposal Package Description gave a list of 8 functionalities important for MPEG-4, that were not supported (or not well-supported) by existing standards: *Content-based multimedia data access tools, Content-based manipulation and bitstream editing, Hybrid natural and synthetic data coding, Improved temporal random access, Improved coding efficiency, Coding of multiple concurrent data streams, Robustness in error-prone environments, and Content-based scalability.*

Proposals for audio and video coding technology were received and evaluated [6, 7], during two MPEG meetings, end 1995 and early 1996. The evaluation included extensive subjective testing for some functionalities [6, 8]. After that, collaboration on the technical work towards defining the standard began. For the work on coding of synthetic material, carried out in the SNHC group, there was an additional call for proposals [9] and the proposals received were evaluated during the September/October 1996 MPEG meeting in Chicago. (SNHC stands for Synthetic and Natural Hybrid Coding.) More calls may follow if at any point in the MPEG-4 development process specific needs are identified for which no technology is available.

In the collaborative phase, 'Verification Models' (VM's) are drafted for each part of the standard, and these are improved during every meeting. Verification Models currently exist for Systems, Audio, Video and SNHC, and they are maintained by the respective subgroups 10, 11, 12, 13. The Systems, Video and Audio VM's will eventually evolve to become parts 1, 2 and 3 of the MPEG-4 standard, while the SNHC VM will be integrated in the Audio and Video parts. The reasoning behind this integration is that sometimes it will be very difficult to make a distinction between what is natural and what is synthetic. For example the texture coding used in natural images can also be used to code the textures to be mapped in synthetic models. The same applies for an audio coding scheme that transmits phonemes: the phonemes could be either

computer-generated, or be the result of an encoder performing an analysis of real speech.

In July 1997, a second round of evaluation will be organized 15. This round of evaluation will be focused on the verification of the MPEG-4 work done so far. On the one hand it includes comparing the Audio and Video Verification Models to both existing standards as well as to new, emerging, technology, and on the other hand it comprises checking the status of the VM's against the requirements for the standard 16. The remaining time schedule for the development of the MPEG-4 standard looks as follows:

November 1996 Working Draft (WD)  
November 1997 Committee Draft (CD)  
July 1998 Draft International Standard (DIS)  
November 1998 International Standard (IS)

Although this time schedule looks aggressive, there is a compelling need to keep to these deadlines. Many of the problems that MPEG-4 tackles urgently need a solution, and if this solution is not provided by MPEG, other (partial) solutions will emerge - proprietary solutions, at the cost of interworking.

### 1. About Profiles and Levels

While the technical development of the standard continues, MPEG's Requirements group is working on the definition of MPEG-4 'profiles' 17. Profiles, as in MPEG-2, are a choice of tools from the complete toolbox, addressing a cluster of functionalities that can be used to enable one or more classes of applications. It is expected that separate Systems, Video and Audio profiles will be defined. Profiles can exist at certain levels, which means that parameters and constraints are fixed to satisfy these classes of applications. The objective is to keep the amount of profiles small, and to enable as many application areas as possible with a single profile. In general, the policy is as follows:

1. A profile will enable applications based on their shared need for similar tools. When requirements for a new application area are brought to MPEG, an attempt is made to fit the new application area in an existing profile, if necessary by extending an existing profile to include more tools;
2. If extending the existing profiles would burden the other applications already supported by the profiles too much, a new profile can be created;
3. Also when all existing profiles are too 'heavy' for the new application's requirements, a new profile can be created.

Profiles in MPEG-4 have two main objectives. First, they define 'conformance points', needed to determine compliance with the toolbox standard that MPEG-4 will be. The objectives are to ease interoperability and decrease complexity. Second, they drive the technical work, as functionalities that are not used in any profile are deemed unnecessary, and work on them will be eventually abandoned. The reverse may also happen: in the definition of the profiles, new requirements could appear. The definition of the profiles is driven by companies working in MPEG, that want to use the new standard for building applications and offering services; this ensures that the technical work carried out in MPEG makes business sense.

The first profile that was defined, to enable 'Real-Time Communications', was drawn up based on requirements provided by the ITU-T Study Group 15 Low Bitrate Coding Group. The design of a second profile, meant to allow 'Content-Based Storage and Retrieval', was started during the Chicago meeting, in October/November 1996, and a third profile, addressing multimedia broadcast applications, has been proposed in November 1996, at the Maceió meeting 17. In MPEG-2, the profiles were organized in a very hierarchical way, each new profile being a superset of all the previous ones. It is unlikely that MPEG-4 can be organized in such a strictly hierarchical way, because there are many more degrees of freedom when compared to MPEG-2.

### 1. Concluding Remarks

While previous audiovisual representation standards were basically limited to translating analog signals into the corresponding digital versions, (e.g. analog vs. digital videotelephony, analog vs. digital TV) MPEG-4 is the first intrinsically digital audiovisual representation standard in the sense that it organizes the scene as a composition of arbitrarily shaped, discrete pieces rather than rectangles of digitized data samples, with corresponding audio.

In its way of defining a coded representation and design of the systems layer, MPEG-4 seeks to technically unify previously distinct service models coming from the historically separate worlds of Information Technology (IT), Telecommunications and Broadcast.

As the first true multimedia data representation standard, MPEG-4 allows the integration of different types of data, with a representation that is optimized for each data type. Moreover, the user will be given the power to interact with the content and control over the way the information is presented. Finally, the communication can include flows to and from different

sources simultaneously, over a heterogeneous channel that should ideally be transparent to the user.

Summarizing all in one sentence: the audiovisual representation standard MPEG-4 establishes a flexible environment, that can be customized for specific applications, provides content-based interactive capabilities with most data types and over most types of available networks, and can be adapted in the future to take advantage of new developments in coding technology.

## 1. Acknowledgments

R. Koenen and F. Pereira acknowledge the support of the European Commission of their work on the MPEG-4 standard, under the RACE program - project MAVT R2072, and the ACTS program - project MoMuSys AC098. F. Pereira acknowledges the support of Junta Nacional de Investigação Científica e Tecnológica for the support of this work under the project 'Processamento Digital de Áudio e Vídeo'.

## 1. References

- [1] L. Chiariglione (Convenor), "MPEG-4 project description", Document ISO/IEC JTC1/SC29/WG11 N1177, Munich MPEG meeting, January 1996
- [2] ITU-T Recommendation H.324, "Terminal for low bitrate multimedia communication", 1996
- [3] MPEG Systems Group, "Systems Working Draft, version 2.0", Document ISO/IEC JTC1/SC29/WG11 N1483, Maceió MPEG meeting, November 1996
- [4] L. Chiariglione (Convenor), "MPEG-4 call for proposals", Document ISO/IEC JTC1/SC29/WG11 N820, Singapore MPEG meeting, November 1994
- [5] MPEG AOE Group, "Proposal Package Description (PPD) - Revision 3", Document ISO/IEC JTC1/SC29/WG11 N998, Tokyo MPEG meeting, July 1995
- [6] MPEG AOE Group, "MPEG-4 testing and evaluation procedures document", Document ISO/IEC JTC1/SC29/WG11 N999, Tokyo MPEG meeting, July 1995
- [7] Ad hoc Group on MPEG-4 Video Testing Logistics, Document ISO/IEC JTC1/SC29/WG11 MPEG95/532, "Report of the ad hoc group on MPEG-4 video testing logistics", Dallas MPEG meeting, November 1995
- [8] T. Alpert, V. Baroncini, D. Choi, L. Contin, R. Koenen, F. Pereira, H. Peterson, "Subjective evaluation of MPEG-4 video codec proposals: methodological approach and test procedures", in this issue
- [9] MPEG Integration Group, "MPEG4 SNHC call for proposals", Document ISO/IEC JTC1/SC29/WG11 N1195, Florence MPEG meeting, March 1996
- [10] MPEG Systems Group, "MSDL VM", Document ISO/IEC JTC1/SC29/WG11 N1484, Maceió MPEG meeting, November 1996
- [11] MPEG Audio Group, "MPEG-4 audio video verification model 2.0", Document ISO/IEC JTC1/SC29/WG11 N1503, Maceió MPEG meeting, November 1996
- [12] MPEG Video Group, "MPEG-4 video verification model 5.0", Document ISO/IEC JTC1/SC29/WG11 N1469, Maceió MPEG meeting, November 1996
- [13] MPEG SNHC Group, "Draft specification of SNHC verification model 2.0", Document ISO/IEC JTC1/SC29/WG11 N1454, Maceió MPEG meeting, November 1996
- [14] L. Chiariglione (Convenor), "Verification model (VM) development and core experiments", Document ISO/IEC JTC1/SC29/WG11 N1375, Chicago MPEG meeting, October 1996
- [15] MPEG Test Group, "Test methods and procedures for July'97 MPEG-4 tests", Document ISO/IEC JTC1/SC29/WG11 N1442, Maceió MPEG meeting, November 1996
- [16] MPEG Requirements Group, "MPEG-4 requirements", Document ISO/IEC JTC1/SC29/WG11 N1495, Maceió MPEG



meeting, November 1996

[17] MPEG Requirements Group, "MPEG-4 profiles", Document ISO/IEC JTC1/SC29/WG11 N1494, Maceió MPEG meeting, November 1996